Using AI to Remove Moiré Patterns and Aliasing in Point Cloud Renders

Group 34: **Brian Formento, Callum Anderson, Damyan Kostadinov, Douglas Morgan, Harvey Brown.**
Supervisor: **Dr. Sasan Mahmoodi**
Second Examiner: **Dr Stuart Middleton, Dr. Tim Norman**

UNIVERSITY OF Southampton

Introduction

ScanLAB Projects is a technology company specializing in creating fly-through videos of real-world environments. LIDAR scanners are used to produce a set of equirectangular depth-maps, which are combined into a single point-cloud of the scene. These point clouds are used in the place of traditional mesh based models to render CGI videos in a traditional rasterizer. This means surfaces often appear semi-transparent - a signature look which distinguishes ScanLAB Project's work.

This rasterization of overlapping point cloud renders into frames of a video produce visual artifacting that makes the videos unfit for clients use such as:

- **Aliasing:** Distortion manifesting as very high frequency striping in areas of the image where flat textures should be seen.
- **Moiré Patterns:** A key challenge for automation as the Moiré patterns present are very tightly packed compared to other aliasing in the images.
- **Temporal Instability:** Differences in noise from one frame to another cause flashing, 'white-noise' like artifacting in the resulting video.

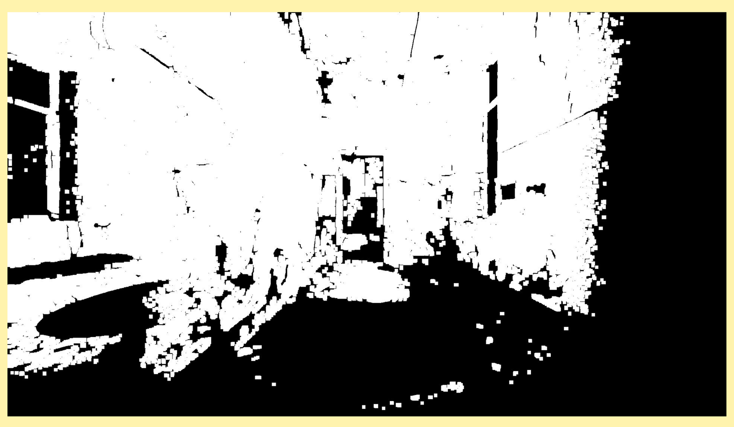
These issues are currently hand fixed in a frame by frame manner by masking sections of each frame and applying blurring.

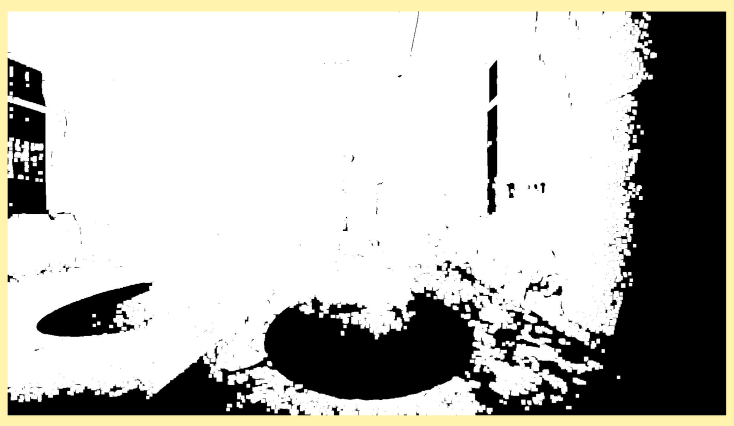
Metrics for Success

Four metrics were employed at various stages to quantify the success of our solutions - typically this is done by visual inspection, but this is a poor comparison of different techniques.

- **Mean Squared Error/Peak Signal to Noise Ratio:** Average squared difference between two images. PSNR is inversely proportional to MSE and normalises for image intensity scaling[3].
- **Structural Similarity Index:** Measures similarity of luminances, contrast and structures to assess perceived quality.[1][6]
- **Strict Pixel Difference:** Absolute difference in RGB value of each pixel in two images.
- **Laplacian Filter:** A measure of the second spatial derivative of an image used for edge detection and detecting blurriness.

Image Morphology





We experimented with using binary opening & closing operations to construct noise masks of the images, which could then be used to guide gaussian filters or L_R deconvolution. We found that most thresholding algorithms[4] included too much of the image in the 'background', and so we used a modified Otsu's method. We performed binary opening on these images with a 9x9 template to produce the binary masks seen above. The original image then had a gaussian filter applied over the dark areas of the mask to produce the final reduced noise version (not pictured).

Grayscale Morphology

You can also perform opening & closing operations on grayscale images. Using the same binary masks (above) as a guide, we experimented with replacing gaussian blurs with grayscale closings of various sizes. This was much better at removing the structural noise, but also introduced unsightly blotches to the image.

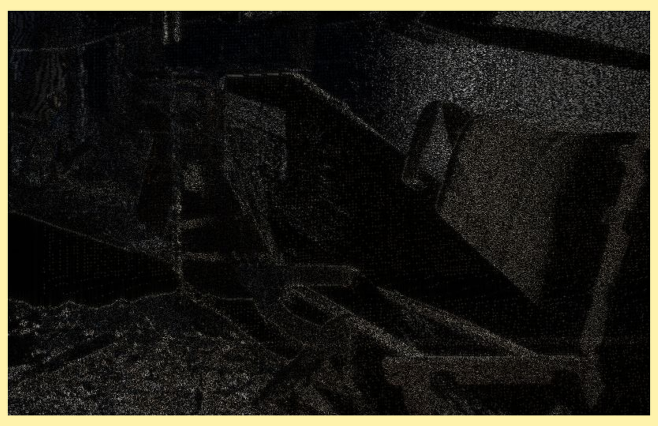
Lucy-Richardson Deconvolution

Out of all the approaches we used, LR[2] deconvolution gives best results in terms of both Similarity Index and level of blurriness. Below are (from left to right) original image, hand fixed image and output from the Lucy-Richardson.

We can see that the noise is removed to a satisfactory level and the image is relatively sharp. Lucy-Richardson deconvolution was the best analytical approach we tried.

Fourier Transforms

Using Fourier transforms we can identify the areas of suspicious co-localisations (noise) within an image[5] and take an extract of that image that represents only the top percentage of frequencies. In this case we take the top 35% of frequencies and get an accurate mapping of the noise which needs to be removed.



Fourier Masking

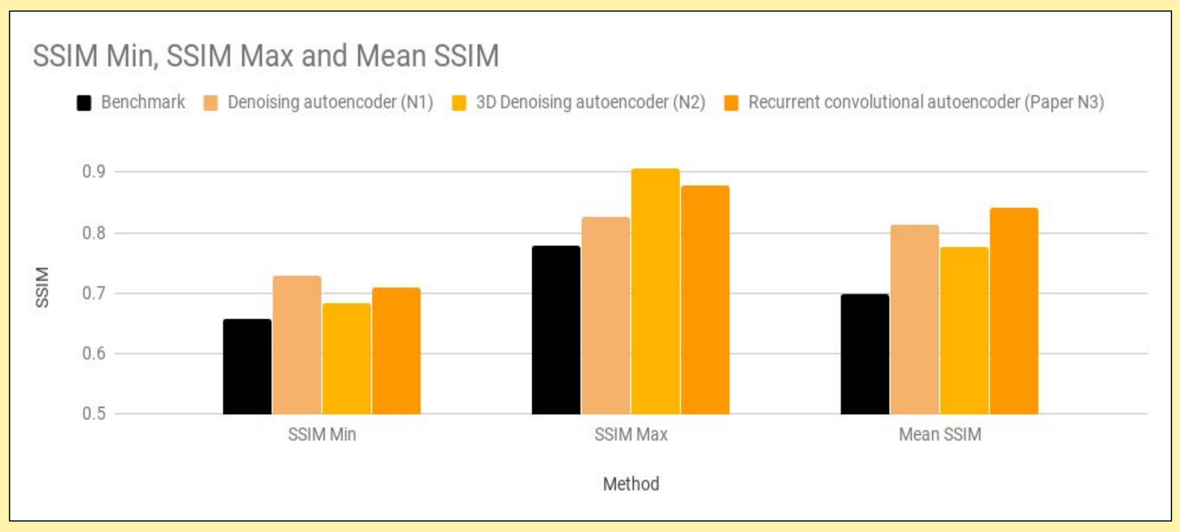
Using fourier transforms as non-binary layer masks over the original image, we show how fine detail in non-noisy areas of the original images, may be restored within a Lucy-Richardson Deconvolution. Which can over blur some of the smaller details of the image where noise removal is not required.

Locations of the masking

LR masked over the raw image

Deep Learning

Three autoencoders have been engineered and tested against the Benchmark. All solutions show a degree of improvement.



Such methods, however, require more resources than the alternatives. Towards the later stages of the project the 3D autoencoder achieved above 0.9 SSIM, but because the wrong training/testing procedure (sliding window) was used blending between images occurred. Individual images, although blended, performed well on the SSIM score and visually looked well reconstructed. Given that deep learning is relatively new to the alternatives, this gave the team good confidence that an optimized solution with the right procedure could, in theory, outperform analytical methods.

Conclusion

Performance is hard to *quantify* when determining image *quality*, but we have attempted to find reliable metrics to guide our research. Inspection by a human remains the true final measure of success however. We believe that one of, or a combination of, the techniques demonstrated will be sufficient for the automatic removal of the noise, although they are not perfect and could all do with additional research.

We find that there is a clear distinction between what we refer to as the *analytical* methods (deconvolution, fourier analysis, and morphology) and the *deep learning* methods (autoencoders). We have found the analytical methods far easier to implement and, crucially, far more consistent in output (in their current state).

As such, we conclude that the best current method for automatically removing the kinds of noise present in ScanLAB Project's videos would be whichever of the analytical solutions they deem to produce the highest quality images.

References

[1] C. R. A. Chaitanya, A. S. Kaplanyan, C. Schied, M. Salvi, A. Lefohn, D. Nowrouzezahrai, and T. Aila, "Interactive reconstruction of monte carlo image sequences using a recurrent denoising autoencoder", ACM Transactions on Graphics, vol. 36, no. 4, pp. 1–12, 2017. doi: 10.1145/3072959.3073601.

[2] W. H. Richardson, "Bayesian-based iterative method of image restoration*", Journal of the Optical Society of America., 1972. doi: 10.1364/JOSA.62.000055.

[3] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? a new look at signal fidelity measures", IEEE Signal Processing Magazine, vol. 26, no. 1, pp. 98–117, 2009, issn: 1053-5888. doi: 10.1109/MSP.2008.930649.

[4] N. Otsu, "A threshold selection method from gray-level histograms", 1979.

[5] B. Coulange and L. Moisan, "An aliasing detection algorithm based on suspicious colocalizations of fourier coefficients", Oct. 2010, pp. 2013 –2016. doi: 10.1109/ICIP.2010.5651195.

[6] Z Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity", IEEE Trans. Image Process., vol. 13, pp. 600–612, Jan. 2014.